A Methodological Comparison of Giving Surveys: Indiana as a Test Case

Patrick M. Rooney, Kathryn S. Steinberg and Paul G. Schervish

Nonprofit and Voluntary Sector Quarterly 2001; 30; 551
DOI: 10.1177/0899764001303011

A Methodological Comparison of Giving Surveys: Indiana as a Test Case

Patrick M. Rooney
Kathryn S. Steinberg
Indiana University–Purdue University Indianapolis

Paul G. Schervish
Boston College

Every 4 years, the Center on Philanthropy at Indiana University conducts a telephone survey (called Indiana Gives) of the giving and volunteering behaviors of Indiana citizens. In preparing to conduct Indiana Gives for 2000, a larger methodological question was asked: How much does survey methodology matter in generating accurate measures of giving and volunteering? In this most recent wave of the Indiana survey, conducted in October and November 2000, eight groups of approximately 100 randomly selected Indiana residents were asked to complete one of eight surveys related to giving and volunteering. It was found that the longer the module and the more detailed its prompts, the more likely a household was to recall making any charitable contribution and the higher the average level of its giving. These differences persisted even after controlling for differences in age, educational attainment, income, household status, race, and gender.

There is growing value and interest in measuring philanthropy at both the national and local levels. This is amply demonstrated by recent research at the state and national levels, such as studies on giving in California (O'Neill, 2000; O'Neill & Roberts, 2000) and Michigan (Council of Michigan Foundations & Michigan Nonprofit Association, 2000); research by Michael Hall (2000) and the Canadian Centre for Philanthropy (Hall et al., 1998) on giving in Canada; research by Arthur Kirsch, Michael McCormack, and Susan Saxon-Harrold

Note: We would like to thank the participants of the Giving Conference and especially Gene Tempel and Kirsten Grønbjerg for their helpful comments on this article. In addition, we are deeply indebted to Brian Vargus and the Public Opinion Laboratory at Indiana University–Purdue University Indianapolis for assistance with the data collection. We would also like to acknowledge William Chin, Mark Wilhelm, Partha Deb, and Richard Steinberg, who assisted us with the regressions.

Nonprofit and Voluntary Sector Quarterly, vol. 30, no. 3, September 2001 551-568

(2001) and Independent Sector (IS) (1999); research by John Havens and Paul Schervish (1996, 2001); *Giving USA* (American Association of Fundraising Counsel Trust for Philanthropy, 2000); and research by the National Center for Charitable Statistics (2000). There is a new annual report on giving in the Netherlands (Schuyt, 1999) and a series of surveys on giving and volunteering in Ireland (Ruddle & Mulvihill, 1995, 1999). A growing number of studies of local areas such as Erie, Pennsylvania (Kurre & Andaleeb, 1998); Kansas City (Greater Kansas City Community Foundation and Affiliated Trusts, 1998); the greater Washington, D.C., region (Washington Regional Association of Grantmakers, 2000); Silicon Valley (Community Foundation Silicon Valley, 1998); and South Florida (The Donors Forum, 1999) confirm that the measurement of philanthropy is a topic of growing interest at the city and regional (substate) levels as well. Kurre and Andaleeb (1998) cited similar studies conducted in Detroit; Orange County, California; and New Hampshire. Although not meant to be an exhaustive list of all these types of studies, this list does demonstrate the increasing prevalence and importance of such work.

Although charity officials and the public are interested in knowing the level and trends of philanthropic behavior, the research community rightly remains vigilant about the validity of the findings generated by survey research and about how to improve this research. Because estimates of the amount of philanthropic behavior appear to rely in such large part on the methods and measures of each survey, this has been a topic of enduring debate within the academic and practitioner communities. For example, in 1993, an entire issue of *Voluntas* was devoted to this topic ("The Dimensions," 1993). *The Chronicle on Philanthropy* published several articles (e.g., Blum, 1998; Lipman, 1999) that were critical of the methodology used in *Giving USA*, to which the American Association of Fundraising Counsel Trust for Philanthropy responded (Raybin & Kaplan, 1999), debating some of the points made in the articles. Similarly, Schervish and Havens (1998a) published a critique of the methodology used by IS, to which IS responded (Hodgkinson & Weitzman, 1998). Schervish and Havens (1998b) subsequently replied to IS's response.

The need for further work on the proper research techniques for studying giving and volunteering is reinforced by the fact that, to our knowledge, all the local and regional studies produce results indicating that the rates of giving in their respective regions are above the national averages cited by IS and in *Giving USA*. Certainly, a "Lake Wobegon effect" may be in place. That is, the local regions selected for study may in fact have higher levels of giving, in which case differences in methodology would not cause the higher results. For example, it may be the case that urban areas (which are most frequently studied) are wealthier and more philanthropic than rural areas (which would be included as part of national studies). However, the differences in methodology and these consistently higher giving estimates make one more skeptical of whether the differences are real or attributable to methodological differences. Indeed, previous studies of regional differences in giving have shown that there is wide variation in generosity between different metropolitan areas and

states (Greene, Millar, & Moore, 1994; Wolpert, 1993) because of various factors such as affluence, unemployment, corporate and civic presence, political conservatism, and population stability. The consistently higher findings of studies with more careful methodologies does suggest that the differences may be attributable as much to differences in methodology as to actual differences in giving reported in local and regional studies.

Since 1991, the Center on Philanthropy at Indiana University (IU) has published a study of philanthropy in Indiana on a quadrennial basis. As we prepared to revise this study last year, we asked a larger methodological question: How much does survey methodology matter in generating accurate measures of giving and volunteering? For instance, to what extent do the following factors forge estimates of individual philanthropy: the number of prompts by the method of fundraising contact (e.g., direct mail, special events, etc.), the number of prompts by the area of donation (e.g., National Taxonomy of Exempt Entities subsectors such as education, environment, etc.), the number of prompts by both the method of fundraising contact and the area of donation, the order of prompts, inducements to participate in surveys, and so forth? Would a simple and very short survey generate comparable results, producing considerable cost savings in studying the sector if the survey yielded results comparable with those of longer surveys?

Our research was significantly informed and enhanced by the seminal survey work done by our predecessors, such as IS (1999), Hall (2000), and O'Neill (2000), as well as the innovative diary study conducted by Havens and Schervish (2001). As such, rather than implying a criticism of this previous work, our current research is in fact a continuation and logical extension of many earlier efforts.

The goals of this project were both empirical and methodological: (a) to update measures of giving and volunteering in Indiana; (b) to systematically compare various methodological survey techniques for estimating giving and volunteering; (c) to assess whether differences in survey methodology such as in the wording, order, number of prompts, and levels of respondent inducements produce different findings; and (d) to discern whether some practical rules of thumb can be formulated for making sense of different survey results and improving future surveys. We are not sure whether it will ever be useful for researchers to settle on the core of a single best survey strategy for estimating giving and volunteering. But, assuming this to be a worthy goal, our current research provides only a modest first step toward figuring out some of what might constitute the elements of such a research design.

METHOD

We constructed a multipronged research design whereby we compared and contrasted the findings of different combinations of survey techniques. We replicated central design elements of surveys done by IS (several prompts

based on giving by "area" of donation, such as education, health care, etc.), Hall (several prompts based on "method" of fundraising contact, such as direct mail, special events, etc.), and an early draft of a survey module designed by the University of Michigan and the Center on Philanthropy at IU for a longitudinal study, the Panel Study of Income Dynamics (PSID) (several prompts based on key areas of giving, with a prompt for method of giving in the introduction). We refer to these respectively as the Area, Method, and PSID modules. In addition, we used a Very Short module ("Did you give last year? If so, how much?" and "Did you volunteer last year? If so, how much?"). We also fielded two modules that enhanced a technique developed in O'Neill and Roberts's (2000) California survey (crossing several prompts for giving by area and then by method of contact). One of these two modules (IU-Method-Area) prompted first by method of fundraising contact, then by subsector of giving for each possible combination. The other module (IU-Area-Method) prompted first about the subsector of giving, then about the method of fundraising contact for each possible combination. There was also a Validity Check module in which the sectors that typically receive the most gifts dollarwise (i.e., religion, education, health, human services, arts, culture, and humanities) were prompted extensively, but other sectors were captured only in a catchall question. In addition, for the two longest modules (which could be as long as 90 minutes each), the telephone researchers were instructed to offer varying levels of inducements randomly ($0, $10, $25, and $50) to ascertain whether this affected the response rates and/or the levels of giving reported. Table 1 summarizes the modules by the types of giving prompts, the number of questions, and inducement levels.

We also included a module that focused on volunteering behaviors. The measurement of volunteering is a separate methodological question from the measurement of giving. Some researchers (e.g., Independent Sector, 1999) have prompted volunteering by subsector; others (e.g., Hall, 2000, 2001) have prompted by method of volunteering (cleaning, serving food, reading, etc.). O'Neill and Roberts's (2000) California survey queried volunteering by subsector, then by method. An additional question concerns whether the measurement of giving should focus on only "formal" volunteering (through organizations) or include "informal" (on one's own) volunteering. Our study included different combinations of these techniques in our modules, but, because of space considerations, these results are not reported here.

To collect our data, the Public Opinion Laboratory at Indiana University–Purdue University Indianapolis used random digit dialing of households to obtain samples using the eight surveys to measure personal philanthropy in Indiana. To ensure comparability in sampling and questioning, the surveys were conducted by small teams whose members were trained across the entire project. The team members were instructed to collect samples of at least 100 respondents in each of eight distinct modules and to monitor the response rates among the various modules. Each respondent was to participate in only

Table 1. Modules by Types of Prompts and Levels of Inducement

	Very Short Module	*IU-PSID Module*	*Validity Check Module*	*Area Module*	*Method Module*	*IU-Area-Method Module*	*IU-Method-AreaModule*	*Volunteer Only Module*
Number of questions	16	137	168	337	368	676	680	223
Prompts	None	Prompt by subsector of giving with short prompt for method	Prompt by major sub-sectors in detail, others in general	Prompt by subsector of contribution	Prompt by method of contact	Prompt by subsector of contribution, then by method of contact	Prompt by method of contact, then by subsector of contri-bution	Prompt about volunteer behavior only
Sample size	110	113	101	106	103	124	124	104
Inducements								
None	110	113	101	106	103	28	29	104
$10						29	30	
$25						33	30	
$50						34	34	

Note: IU = Indiana University, PSID = Panel Study of Income Dynamics.

one of the modules. Because each team worked independently, the actual number of participants surveyed varied somewhat from module to module, ranging from 101 to 124.

It should be noted that in each module, we asked respondents about household giving. This is the method that IS and most U.S. researchers have used when collecting data on giving. Household giving is of interest because it is more comparable with data reported on tax returns and, therefore, could in theory be verified with Internal Revenue Service (IRS) data. IRS returns may be the best memory "triggers" for survey respondents. (Note that IS [1999] has found that the reliability for household volunteering is quite low, so we focused on volunteering done by the respondents.) However, the focus on household giving is noteworthy in the replication of Hall's (2000; Method) module because that survey of giving in Canada focused on individual giving rather than household giving. Although we did try to replicate the prompts used in Hall's study, we queried household giving for comparability across the modules. This discussion raises a more general methodological question that scholars must test in the field. Namely, how is the measurement error affected when surveys ask about individual or household philanthropy? The key considerations are greater knowledge and recall of one's own philanthropy, but many may recall how much they listed as itemized deductions, which would typically be a joint total (even married persons filing separately would be likely to pool their contributions and list them under one person because this would minimize their combined tax burden).

The emphasis in this article is on the measurement of total formal giving, that is, contributions to charitable or nonprofit organizations or groups that are legally deductible on Schedule A income tax forms. We also measured differences across demographic groups: age, income, household status (couples vs. singles and/or marital status), race, educational attainment, and tax status (i.e., itemizers vs. nonitemizers). Three of our modules (Area, IU-Area-Method, and IU-Method-Area) included questions on informal giving, that is, contributions that are not collected through particular organizations or groups. We do not emphasize these results in this article except to raise the methodological question of whether the measurement of giving should focus on only formal giving or also include informal giving.

To test the data on formal giving, we first used ANOVAs to test whether the samples receiving each module had comparable demographic characteristics and reported giving. Then, we performed several multivariate analyses (as suggested by O'Neill, 2001) to see whether differences in mean reported giving across modules could be explained by variations in sample characteristics or appeared to be pure effects of the module administered. To do so, we explained donations in a regression framework by including a set of dummy variables for six modules, along with age, age squared, income (seven categories), itemization status, race (White and all others), marital status (married or cohabitating vs. single, widowed, and divorced), education (high school or less, some college, bachelor's degree, and graduate or professional school),

Table 2. Advantages and Disadvantages of Regression Models

Statistical Method	*Advantage*	*Disadvantage*
Tobit on full sample	• Eliminates selection bias • Does not generate negative predicted donations	• Not robust to nonnormal or heteroskedastic errors • Enforces proportionality between a variable's effect on the probability of giving and the sizes of the donations of those who give
Heckman two-stage	• Eliminates selection bias • Separates a variable's effect on probability of giving from its effect on the amount given	• Not robust to nonnormality
Ordinary least squares on donors only	• Robust to nonnormal errors and heteroskedasticity	• Suffers from selectivity bias
Ordinary least squares on full sample	• Robust to nonnormal errors and heteroskedasticity	• Suffers from truncation bias (assumes a symmetrical distribution, including the possibility of negative gifts)

and gender. If there were pure module effects, they would show up as significant coefficients for the module dummy variables.

Unfortunately, the error terms in these regressions did not obey the classical assumptions that justify the use of ordinary least squares (OLS). Donations cannot be negative, so the error term had a truncated distribution. In addition, giving data appeared to have a nonnormal error structure (e.g., Bradley, Holden, & McClelland, 1999). Under these circumstances, OLS is biased and inconsistent. Because there is no commonly accepted ideal remedy for these problems, we report results from four different approaches (Tobit, Heckman two-stage, OLS on the full sample, and OLS on positive donors only) and hope that a consistent picture emerges. The advantages and disadvantages of each regression model are presented in Table 2.

DATA AND BIVARIATE RESULTS

The data set used in this article includes seven modules, each with samples of between 100 and 125 respondents. Each module includes questions about giving and volunteering. Hence, we have 781 respondents from the seven modules who addressed both the time and treasure components of personal philanthropy and a total of almost 900 respondents who replied to questions about their volunteer efforts (the seven modules in this article and one module that focused exclusively on volunteer behaviors). We start this section by delineating the major differences and similarities between the various

sampling frames or modules to enable readers to make their own judgments. In the next section, we present regression analyses, which allowed us to examine the results while controlling for differences in the methodologies and other factors such as income and educational attainment.

Table 3 presents a comparison of the demographic characteristics of each sample. We discuss briefly only those sample characteristics that are statistically significantly different from the combined or total sample. Compared with the total sample means and proportions, the Area module sample contained significantly fewer couples, more minority and more highly educated participants, more participants in the top income category, and more who itemized deductions and gifts. The Very Short sample included significantly older participants, more White participants, more participants who had never attended college, fewer in the high-income group, and fewer who itemized deductions and gifts. The Method survey sample included significantly more participants with some college education and more who had attended graduate or professional school, as well as more participants in the high-income group. Two samples (the IU-PSID and Validity Check modules) had significantly higher proportions of participants with high school diplomas or less education, although the IU-PSID module also had a higher proportion of participants who itemized donations. There were also two groups (the IU-Method-Area and IU-Area-Method modules) with higher proportions of participants with some college education. The IU-Area-Method sample also had significantly more participants in the low-income group.

An interesting difficulty that we ran into was that there were differences across modules in the reporting of income and other key variables. Probit regression analyses, not reported here, indicated that the module used affected the probability of answering the income question. For example, people were least likely to report their incomes in the Area and Very Short modules and most likely to report their incomes in the IU-Area-Method module. As shown in Table 3, only 61.8% of the respondents reported their household incomes in the Very Short module and 63.2% in the Area module, whereas 79.1% reported their incomes in the IU-Area-Method module. Further research is needed to determine if this problem was particular to this study or whether it applies more generally.

In Table 4, the simple means and medians of formal giving for each of the modules are compared, and some interesting similarities and differences are shown. The means for the Very Short and IU-PSID modules were significantly lower and the mean for the IU-Method-Area module was significantly higher than the total sample mean. The Very Short survey had a mean value of $504 and a median value of $50, both of which were about one half or less of the other modules. The means for the others ranged from $867 (IU-PSID) to $2,336 (IU-Method-Area), and the medians ranged from $115 (Method) to $645 (IU-Method-Area). It is worth noting that the second highest mean and median values came from the IU-Area-Method module, which, along with the IU-Method-Area module, was the longest module tested. An even bigger

Table 3. Demographics

Characteristic	*All Participants*	*Very Short*	*IU-PSID*	*Validity Check*	*Area*	*Method*	*IU-Area-Method*	*IU-Method-Area*
Sample size	781	110	113	101	106	103	124	124
Female (%)	65.0	68.2	68.1	60.4	65.1	63.1	62.9	66.9
Couples (%)	61.7	60.9	59.3	61.0	54.7***	62.2	54.8	60.5
White (%)	86.6	95.5***	86.7	83.8	79.2**	89.2	84.7	88.7
Age								
Mean	47.5	52.6***	48.4	45.8	47.3	46.6	44.57	47.1
Median	46	52	47.5	42	46.5	48	42.5	43
Minimum	18	18	18	18	18	18	18	18
Maximum	94	87	94	86	84	87	89	88
Education (%)								
High school diploma or less	36.4	59.3***	47.3***	48***	40.5	30.4	37.1	37.1
Some college	28.0	29.0	31.3	27.0	22.7	38.2**	35.5*	36.3**
Bachelor's degree	14.0	8.2*	15.0	13.0	23.6***	13.7	18.5	16.9
Graduate or professional school	8.5	3.6*	6.3	12.0	13.2*	17.6***	8.9	9.7
Joint tests		***			**	***		
Income (%)								
$0 to $40,000	36.4	40.9	31.9	35.6	31.1	33.0	44.4*	36.3
$40,000 to $80,000	22.0	17.3	26.5	23.8	18.9	19.4	22.6	25.0
$80,000 or more	10.9	3.6**	10.6	5.9	13.2	18.4*	12.1	12.1
Total reporting income	69.3	61.8	69.0	65.3	63.2	70.8	79.1	73.4
Joint tests		**				*		
Percentage of sample with incomes > $120,000	2.9	2.9	3.8	1.5	11.9***	6.8	2.0	3.2
Percentage who itemized deductions	40.2	34.0**	44.2**	37.6	48.1***	34.7	40.3	34.7
Percentage of itemizers with donations	73.4	55.9	72.3	89.2	76.0	68.1	77.1	73.8
Percentage of sample with itemized gifts	28.7	17.3***	30.1	32.7	35.8*	31.1	29.8	25.0

Note: IU = Indiana University, PSID = Panel Study of Income Dynamics. Asterisks indicate significant differences between individual modules and total sample means or proportions.
*$p < .10$. **$p < .05$. ***$p < .01$.

impact was observed when examining the median values. The median values for the two long modules were over 10 times greater than the Very Short module and 3 or 4 times greater than the Area module. Collectively, these results are prima facie evidence that the number of prompts and the length of the survey do matter in collecting data about individual philanthropy. However, the particular sample using the Very Short survey (i.e., no prompts) also had a lower level of educational attainment and fewer in the high-income group than the other samples, on average. Hence, we will need to further refine this analysis with multivariate analyses.

In the Area, IU-PSID, and Very Short modules, approximately two thirds of the respondents (ranging from 67.9% to 69.6%) made donations, and larger percentages made donations in the other modules. This suggests not that most respondents in the Very Short module forgot that they had made any donations at all but that they might have forgotten some of the donations they had made in the past year or that they just gave smaller gifts than others, on average. The relatively low rates of giving among the relatively short modules suggests that it is likely that at least some of the respondents forgot that they had made any gifts. This is further evidence that more prompts lead to more recall. The longer modules generally yielded higher rates of giving among respondents: Ninety-two percent reported making donations in the Method module, 95% in the IU-Method-Area module, and 98% in the IU Area-Method module.

On the basis of these three important indicators of giving (mean and median dollars given and the percentage of respondents who gave anything), it seems that a greater number of detailed prompts stimulates greater recall. Of course, there is a danger that respondents will report gifts that they did not actually give in an effort to conform to a perceived set of social expectations and/or to please or impress interviewers. Although this is a problem in any type of survey research, it may be exacerbated by repeated questions about giving, which may convey the message to respondents that it is expected or normal to give. Similarly, respondents may be embarrassed or bored if they repeatedly report no giving in surveys that ask about giving by many different areas and/or methods of contacts. These concerns are not readily resolved, but they do indicate the need to verify survey results through independent data sources whenever feasible. However, Havens and Schervish (1996), who used daily prompting in their diary study, found that 100% of their sample made donations during the course of a year. This strongly suggests that fewer prompts lead to an understatement of gifts rather than that more prompts create a false or exaggerated list of gifts.

Another important issue that we investigated in this study relates to refusal rates for the different modules. As one might expect, we found that for both of the modules with incentives (IU-Area-Method and IU-Method-Area), there was a negative correlation (–.84 and –.45 respectively) between incentive amounts and percentage of refusals. Although we do not know the characteristics of people who refused to participate, for those who did participate, there were no statistically significant differences between incentive groups in terms

Table 4. Giving

	All Participants	*Very Short*	*IU-PSID*[a]	*Validity Check*	*Area*	*Method*	*IU-Area-Method*	*IU-Method-Area*
Formal giving ($)								
Mean	1,299	504***	867***	1,051	1,395	1,218	1,462	2,336***
Median	240	50	188	175	141	115	541	645
Percentage who made formal donations	82.4	69.6	68.3	83.0	67.9	92.2	97.6	93.5
Gifts of more than $10,000	19	0	0	3	3	3	3	7
Informal giving ($)								
Mean	NA	NA	NA	NA	2,006	NA	1,375	486
Median	NA	NA	NA	NA	75	NA	150	63
Percentage who made informal donations	NA	NA	NA	NA	69.6	NA	71.4	69.4
Gifts of more than $10,000	NA	NA	NA	NA	5	NA	4	0
Total giving ($)								
Mean	1,299	504	2,044	1,051	3,325	1,218	2,782	2,790
Median	240	50	200	175	538	115	1,104	1,053
Percentage who made some donation	82.4	69.6	65.6	83.0	78.3	92.2	98.4	95.2
Gifts of more than $10,000	19	0	1	3	10	3	12	9
Among contributing households								
Formal giving ($)								
Mean	1,575	724	1,270	1,266	2,053	1,321	1,498	2,497
Median	390	80	450	275	535	140	550	713

Note: IU = Indiana University, PSID = Panel Study of Income Dynamics, NA = not applicable. Asterisks indicate significant differences between individual modules and total sample means or proportions.

a. Subsequent drafts of the PSID survey have obtained higher means in pretests done by the University of Michigan's PSID staff.

***$p < .01$.

Table 5. Refusal Rates

Instrument	*Number of Prompts*	*Incentive*	*Obtained* N	*% of Refusals**
Very Short	16	$0	110	7.90
IU-PSID	137	$0	113	12.10
Validity Check	168	$0	101	10.80
Area	337	$0	106	10.70
Method	368	$0	103	11.30
IU-Area-Method	676	$0	28	18.93
	676	$10	29	20.97
	676	$25	33	19.58
	676	$50	34	11.50
IU-Method-Area	680	$0	29	23.00
	680	$10	30	15.80
	680	$25	31	21.40
	680	$50	34	16.80

Correlations	IU-Method-Area	IU-Area-Method	Total
Incentive and percentage of refusals	–0.45	–0.84	–0.65*
Number of prompts and percentage of refusals (all modules)			0.81***

Note: IU = Indiana University, PSID = Panel Study of Income Dynamics.
*$p < .10$. ***$p < .01$.

of demographics, with the exception that no minority participants received the IU-Method-Area survey with $0 incentives. It is important to note that the incentive amounts did not make a significant difference in the reported amount of total formal giving. Across all modules (including those that did not involve incentives), we found a positive correlation ($r = .81$, $p = .000$) between the number of prompts (which we used as a proxy for length of interview) and refusal rates. Refusal rates ranged from 7.9% for the Very Short module to 23.0% for the IU-Method-Area module with $0 in incentives. These results are summarized in Table 5.

MULTIVARIATE RESULTS

We now turn to the multivariate results, for which we used OLS, Tobit, and the Heckman two-stage regression models to examine the marginal impacts of the independent variables and to test the impacts of differences in the modules more formally. Table 6 presents the main results of our regression analyses. These models assume that the various survey modules did not alter the coefficients of the other variables but that they shifted the regression line up or down depending on the impact a survey module had on reported giving.

Looking at the demographic variables in the regression models (i.e., age, race, gender, education, income, household status, and itemization status), we

Table 6. Regression Results

	Tobit Coefficient[a, c, d, e, g]	*OLS (full sample)* Coefficient[a, c, f, g]	*OLS (donors only) Coefficient*[a, c, f, g]	*Heckman Second Stage (OLS) Coefficient*[a, c, f, g]	*Heckman First Stage (Probit) Coefficient*[b, c, d, e, g]
Constant	–813.92	–57.30	70.15	–589.64	0.10
Age	2.15	–4.14	–7.85	–5.76	0.00
Age squared	0.15	0.25	0.33	0.33	0.00
Income $0 to 20,000	–105.73	–402.06**	–715.47***	–535.41	0.08**
Income $20,000 to 40,000	10.21	–299.84	–578.07**	–380.85	0.10***
Income $40,000 to 60,000	282.65	–65.84	–256.79	78.52	0.18***
Income $60,000 to 80,000	523.38*	356.07	157.86	488.86	0.21***
Income $80,000 to 100,000	983.93***	1,189.61*	930.46	1,290.15	0.281**
Income above $100,000	2,214.66***	3,257.71**	3,059.58**	3,414.07**	0.22***
Itemize	744.26***	920.08***	986.15***	1,124.53***	0.08***
Minority	–161.66	–247.04	–262.03	–314.99	–0.01
Couples	307.09*	258.95	228.34	353.92	0.07**
Some college	505.25***	688.06***	816.26***	883.16***	0.03
Bachelor's degree	514.46**	648.26**	758.03**	840.44**	0.03
Graduate or professional school	639.66**	957.61*	1,002.19*	1,074.42*	0.04
Female	280.62*	144.27	129.26	303.48	0.10
Method module	–491.45*	–657.531	–612.70	–753.57*	–0.12*

(continued)

Table 6 Continued

	Tobit *Coefficient*[a, c, d, e, g]	*OLS (full sample)* Coefficient[a, c, f, g]	*OLS (donors only)* *Coefficient*[a, c, f, g]	*Heckman Second Stage (OLS)* *Coefficient*[a, c, f, g]	*Heckman First Stage (Probit)* *Coefficient*[b, c, d, e, g]
Area module	–703.28**	–310.99	–70.50	–596.27	–0.31***
IU-Method-Area module	378.20	753.47*	815.51*	730.54	–0.09
IU-PSID module	–1,059.05***	–748.61***	–828.60***	–1,341.97	–0.33***
Validity Check module	–365.30	–265.96	–208.60	–467.70	–0.19***
Very Short module	–952.47***	–784.27***	–782.26***	–1,265.20*	–0.30***
Inverse Mills ratio				1270.98	
Log-likelihood	–6,082.82	–7,198.00	–5,986.30	–5,985.92	–265.85
Adjusted R^2	NA	.166686	.156636	.156265	NA
N	763	763	629	763	763

Note: OLS = ordinary least squares, IU = Indiana University, PSID = Panel Study of Income Dynamics.
a. Estimates the change in donation amount due to changes in each variable.
b. Estimates the marginal probability of donating due to changes in each variable.
c. All coefficients for categorical variables are relative to the values of the excluded category for that variable (income unknown, White, single, high school education or less, male, and IU-Area-Method module).
d. Statistical significance is determined for the coefficients on the latent index for donations.
e. Statistical significance with respect to the latent indicator variable.
f. Statistical significance was determined using White's robust standard errors.
g. Table with standard errors or *t* scores is available on request from the authors.
*$p < .10$. **$p < .05$. ***$p < .01$.

found nothing surprising or out of line with other studies. For example, the Tobit analysis indicated that respondents with some college education reported household donations of $295.56 more than respondents with high school educations or less (the excluded category). Thus, a consistent picture emerged across all models: Giving increases with increased income and education and with itemization of taxes. Similarly, race, age, and gender were not significant in explaining differences in giving across the various modules. The lack of significance of the inverse Mills ratio in the Heckman two-stage model suggests that selection bias was not very important in this data set.

Of more interest to this methodological study is the analysis of the dummy variables for different modules. A rank ordering of the module coefficients in each regression model presents an interesting picture. In each regression, the coefficient for the IU-Method-Area module was positive relative to the IU-Area-Method module (which was the excluded category). The coefficients for all other modules were negative, indicating that these modules were associated with lower levels of reported average giving after controlling for income, education, race, household status, age, and so forth. In addition, the two shortest modules (IU-PSID and Very Short) had the largest negative impacts on reported giving in all regressions, which suggests that survey length and the number of prompts in the surveys do matter in terms of gathering data on giving.

To examine whether various modules had impacts on the probability of whether individuals or households gave anything, we looked at the probit results from the Heckman two-stage model. The probit results indicate that holding everything else constant, the coefficients of all modules were negative relative to the IU-Area-Method module, and the coefficients of all but the IU-Method-Area module were statistically significant. That is, controlling for age, income, education, itemization status, race, household status, and gender, the probability of reporting any donations was highest with the IU-Area-Method module. In contrast, the other columns of Table 6 show that the amount of giving, for those who did give, was estimated to be the highest with the IU-Method-Area module. Taken together, these two results once again point to the conclusion that survey length and the number of detailed prompts do matter in estimating household giving.

CONCLUSIONS

Our results show that whether using simple means or multivariate analyses, the longer and more detailed a module, the more likely a household was to make a charitable contribution and the higher the average level of its giving, even after controlling for differences in age, educational attainment, income, household status, race, and gender. Further work is needed to determine whether these results are sustained with larger samples and whether the results from Indiana are representative of the nation. It would also be useful to

conduct research on what kinds of giving people tend to remember and forget without being prompted.

Additional research is also needed to ascertain whether the Very Short survey serves as a useful proxy for some portion of total giving. We would urge caution in using short surveys to estimate household giving until further research can be conducted because there may be regional, racial, social class, or other differences in how respondents react to these surveys. If we could understand in detail what types of underreporting occur in different populations when confronted with short surveys, we could make educated guesses about the true levels of giving.

There is also a need for further research on other important methodological questions. For example, what is the appropriate unit of analysis, household or individual giving? Does increasing the number of prompts increase recall or lead to overestimates of giving (resulting from the desire to conform to expectations or to please interviewers)? In addition, there is a need for further analysis to find the proper remedy for violations of statistical assumptions in giving data.

There remains a need for longitudinal studies to track how changes in micro-level household characteristics over time (e.g., behavior and/or observations as a youth, changes in socioeconomic status, changes in public policy, etc.) affect changes in giving. Also, given that all of the recent local and regional studies of giving report giving levels that are above the national average, there is a need to use the same methodology at the same time across a number of regions to assess whether giving differs by region compared with national averages.

It is unlikely that a telephone survey can adequately sample the most wealthy, yet this group accounts for a relatively large percentage of total giving. According to Havens and Schervish's (personal communication, March 2001) composite analysis of the 1998 Survey of Consumer Finances, approximately 40% of charitable giving is contributed by the 5% of households with incomes above $150,000. Hence, it would be very helpful to have a thorough study of giving among elites compared with everyone else. Such data might allow scholars to supplement what has been learned from other studies (e.g., Auten, Clotfelter, & Schmalbeck, 1997; Schervish & Havens, 2001) to account for giving by elites that is typically missed in random samples. Similarly, it is important to investigate whether there are significant differences in understanding and responding to surveys among various ethnic and racial groups and/or between genders. Most important is a need for survey research to be validated or cross-checked by other means such as IRS data at both the individual and aggregate levels, which assumes that IRS data are themselves accurate. This would increase the credibility of all such work and/or suggest appropriate adjustments and caveats for scholars in this area.

In the end, the research reported here, coupled with the findings cited at the beginning of this article, suggests that when it comes to estimating charitable

giving, methodology is destiny. The more individuals are prompted in detail, the more likely they are to recall having given any gifts, and the higher the total amount of giving they report. Although there is no way to prove this conclusion, the weight of the findings summarized here suggests that increased attention to survey methodology will bear much fruit in identifying more accurately the level and trends of charitable giving in particular and social engagement in general.

References

American Association of Fundraising Counsel Trust for Philanthropy. (2000). *Giving USA: The annual report on philanthropy for the year 1999*. New York: Author.

Auten, G. E., Clotfelter, C. T., & Schmalbeck, R. L. (1997). *Taxes and philanthropy among the wealthy* (Working Paper). Ann Arbor: University of Michigan, Office of Tax Policy Research.

Blum, D. (1998). "Giving USA" revises donation figures downward to correct past flaws. *The Chronicle of Philanthropy*, *10*(16), 30.

Bradley, R., Holden, S., & McClelland, R. (1999). *A robust estimation of the effects of taxation on charitable contributions* (Working Paper). Washington, DC: Bureau of Labor Statistics.

Community Foundation Silicon Valley. (1998). *Giving back, the Silicon Valley way: The culture of giving and volunteerism in Silicon Valley*. San Jose, CA: Author.

Council of Michigan Foundations, & Michigan Nonprofit Association. (2000). *Michigan: Giving and volunteering: Key findings from a 1999 survey of Michigan residents*. Grand Haven and East Lansing: Authors.

The dimensions of the US non-profit sector: A review symposium on the *Nonprofit Almanac* [Special issue]. (1993). *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, *4*(2).

The Donors Forum. (1999). *South Florida gives: A comprehensive report on the state of philanthropy in Palm Beach, Broward, Miami-Dade, and Monroe Counties*. Miami, Florida: Author.

Greater Kansas City Community Foundation and Affiliated Trusts. (1998). *1998 Report on Charitable Giving for Greater Kansas City*. Kansas City, MO: Author.

Greene, E., Millar, B., & Moore, J. (1994). The Midwest's charitable advantage. *The Chronicle of Philanthropy*, *6*(9), 22-26.

Hall, M. (2000). *The measurement of charitable giving: Approaches employed in Canada's National Survey of Giving, Volunteering and Participating*. Toronto, Canada: Canadian Centre for Philanthropy.

Hall, M., Knighton, T., Reed, P., Bussière, P., McRae, D., & Bowen, P. (1998). *Caring Canadians, involved Canadians: Highlights from the 1997 National Survey of Giving, Volunteering and Participating*. Ottawa: Statistics Canada.

Hall, M. H. (2001). Measurement issues in surveys of giving and volunteering and strategies applied in the design of Canada's National Survey of Giving, Volunteering and Participating. *Nonprofit and Voluntary Sector Quarterly*, *30*, 515-526.

Havens, J. J., & Schervish, P. G. (1997). *Our daily bread: Findings from the first diary study on giving and receiving care*. Chestnut Hill, MA: Boston College, Social Welfare Research Institute.

Havens, J. J., & Schervish, P. G. (2001). The methods and metrics of the Boston Area Diary Study. *Nonprofit and Voluntary Sector Quarterly*, *30*, 527-550.

Hodgkinson, V. A., & Weitzman, M. S. (1998). Responding to factual errors present in the Schervish and Havens article, "Embarking on a republic of benevolence? New survey findings on charitable giving," and other comments. *Nonprofit and Voluntary Sector Quarterly*, *27*, 522-528.

Independent Sector. (1999). *Giving and volunteering in the United States, 1999 edition*. Washington, DC: Author.

Kirsch, A. D., McCormack, M. T., & Saxon-Harrold, S.K.E. (2001). Evaluation of differences in giving and volunteering data collected by in-home and telephone interviewing. *Nonprofit and Voluntary Sector Quarterly*, *30*, 495-504.

Kurre, J. A., & Andaleeb, S. S. (1998, November). *Measuring philanthropy in a local area*. Paper presented at the 27th Annual Conference of the Association for Research on Nonprofit Organizations and Voluntary Action, Seattle, WA.

Lipman, H. (1999). Report's numbers are no true measure of charity, critics say. *The Chronicle of Philanthropy*, *11*(16), 27, 30.

National Center for Charitable Statistics. (2000). *NCCS Relative Giving Effort Index: Comparing charitable giving by state*. Washington, DC: Author.

O'Neill, M. (2001). Research on giving and volunteering: Methodological considerations. *Nonprofit and Voluntary Sector Quarterly*, *30*, 505-514.

O'Neill, M., & Roberts, W. L. (2000). *Giving and volunteering in California*. San Francisco: University of San Francisco, Institute for Nonprofit Organization Management.

Raybin, N. L., & Kaplan, A. E. (1999). The key to better research on charitable giving: More financial support. *The Chronicle of Philanthropy*, *11*(17), 53.

Ruddle, H., & Mulvihill, R. (1995). *Reaching out: Charitable giving and volunteering in the Republic of Ireland: The 1994 survey*. Dublin: National College of Ireland.

Ruddle, H., & Mulvihill, R. (1999). *Reaching out: Charitable giving and volunteering in the Republic of Ireland: The 1997/98 surveys*. Dublin: National College of Ireland.

Schervish, P. G., & Havens, J. J. (1998a). Embarking on a republic of benevolence? New survey findings on charitable giving. *Nonprofit and Voluntary Sector Quarterly*, *27*, 237-242.

Schervish, P. G., & Havens, J. J. (1998b). Reply to Hodgkinson and Weitzman. *Nonprofit and Voluntary Sector Quarterly*, *27*, 529-531.

Schervish, P. G., & Havens, J. J. (2001). Wealth and the commonwealth: New findings on wherewithal and philanthropy. *Nonprofit and Voluntary Sector Quarterly*, *30*, 5-25.

Schuyt, T.N.M. (1999). *Giving in the Netherlands 1999*. Houten, the Netherlands: Bohn Stafleu Van Loghum.

Washington Regional Association of Grantmakers. (2000). *A region of givers: The report on philanthropy in the greater Washington, DC region, 2000*. Washington, DC: Author.

Wolpert, J. (1993). *Patterns of generosity in America*. New York: The Twentieth Century Fund Press.

Patrick M. Rooney is the chief operating officer and the director of research at the Center on Philanthropy at Indiana University. He is an associate professor of economics at Indiana University–Purdue University Indianapolis and a member of the Indiana University Philanthropic Studies faculty. He directs the research activities at the Center on Philanthropy, including the Philanthropoy Module for the Panel Study of Income Dynamics (in collaboration with the University of Michigan), the Fundraising and Administrative Cost Study (in collaboration with the Urban Institute), and research for Giving USA.

Kathryn S. Steinberg serves as the assistant director of research at the Center on Philanthropy at Indiana University. She holds a doctorate in educational psychology from Purdue University and has taught at the University of Nebraska at Kearney, the University of Puerto Rico, and Henderson Community College in Kentucky. At the Center on Philanthropy, she assists with a number of research projects, including Indiana Gives, the Philanthropic Giving Index, the Million Dollar List, and four research grant programs.

Paul G. Schervish is a professor of sociology and director of the Social Welfare Research Institute at Boston College. He and John J. Havens have coauthored the 2001 Bankers Trust Study on Wealth and Responsibility and Millionaires and the Millennium, a microsimulation projection of the forthcoming transfer of wealth. He and Mary A. O'Herlihy and John J. Havens have coauthored the 2001 High Tech Donor Study. During the 1999-2000 academic year, Schervish served as distinguished visiting professor at the Center on Philanthropy at Indiana University.